LEGIBILITY NOTICE

A major purpose of the Technical Information Center is to provide the broadest dissemination possible of information contained in DOE's Research and Development Reports to business, industry, the academic community, and federal, state and local governments.

Although portions of this report are not reproducible, it is being made available in microfiche to facilitate the availability of those parts of the document which are legible.

3

LA-UR -88-55

CONF-871120--85

LA-UR--88-55

DE88 005374

TITLE: NOTES ON THE MUF-D STATISTIC

AUTHOR(S): R. R. Picard

SUBMITTED TO ANS Topical Conference, The 3rd International Conference on Facility Operations Safeguards Interface, San Diego, November 29-December 4, 1987

Los Alamos

Los Alamos National Laboratory
Los Alamos, New Mexico 87545

NOTES ON THE MUF-D STATISTIC

R. R. Picard
Los Alamos National Laboratory
P. O. Box 1663, MS E541
Los Alamos, NM 87545
(505) 667-5736

ABSTRACT

Many aspects of the MUF-D statistic, used for verification of accountability data, have been examined in the safeguards literature. In this paper, basic MUF-D results are extended to more general environments than are usually considered. These environments include arbitrary measurement error structures, various sampling regimes that could be imposed by the inspectorate, and the attributes/variables framework.

I. INTRODUCTION

Verification of an inventory or of a reported MUF calls for the remeasurement of a sample of items by an inspectorate followed by comparison of the inspectorate's data with the facility's reported values. Such a comparison is intended to protect against falsification of accounting data that could conceal material loss. In the international arena, the observed discrepancies between the inspectorate's data and the reported data are quantified using the D statistic (see, for example, the IAEA Technical Manual[1]).

Under very general conditions regarding the facility's/inspectorate's measurement error procedures and the inspectorate's sampling regime, the variance of the MUF-D statistic decomposes into three components. The inspection's sensitivity against various falsification scenarios can be traced to one or more of these components. Obvious implications exist for the planning of effective inspections, particularly in the area of resource optimization.

In subsequent sections, properties of the conventional D and MUF-D statistics are examined in detail. Results are then extended to a generalized D statistic, which is useful when two or more types of inspectorate measurements are involved (e.g., the attributes/variables framework). The related contexts go beyond those discussed in the safeguards literature.

II. THE CONVENTIONAL FRAMEWORK

A. Characterization of MUF and D

Derivation of the main results necessarily entails some mathematics. In what follows, matrix algebra is used to develop the theory. Though perhaps unfamiliar to those with nonstatistical backgrounds, the associated notation provides for concise expression of very general formulae. Relation of matrix algebra to the more commonly used (and less readable) summation signs and subscripts is reviewed in many texts[2,3] on linear models. Readers uninterested in such details may ignore them and focus attention on the text.

Let the facility operator's accountability values used in the MUF calculation be represented by the N-component vector o, where o is written in partitioned form as

$$o^T = [o(BI)^T \mid o(R)^T \mid o(S)^T \mid o(EI)^T] \qquad (1)$$

and

- $o(BI)$ is the vector of the $N(BI)$ accountability values corresponding to items in beginning inventory,
- $o(R)$ is the vector of the $N(R)$ accountability values corresponding to items in receipt,
- $o(S)$ is the vector of the $N(S)$ accountability values corresponding to items in shipment,
- $o(EI)$ is the vector of the $N(EI)$ accountability values corresponding to items in ending inventory, and
- $N = N(BI) + N(R) + N(S) + N(EI)$ is the total number of items involved (some items may be counted twice, as would be the case for a given item in both beginning and ending inventory),

where the symbol "T" denotes transposition. Also, let

$$\Sigma_o = \text{the covariance matrix of } o \ . \tag{2}$$

The form of Σ_o can be somewhat complicated, depending on how the facility measures various items, e.g., how accountability weights, concentrations, and so on are measured by available instrumentation.

Let the vector f correspond to o and denote the vector of falsifications. If the facility does not falsify any accountability values, then f is equal to a vector of zeroes. Note that the effects of innocent causes, such as clerical errors, are indistinguishable from those of malevolent falsification and can produce nonzero values of f. In any event, the data reported by the facility is o + f, and it follows that the reported MUF is

$$\begin{aligned} \text{MUF} &= z^T (o + f) \\ &= \text{MUF}_u + F \ , \end{aligned} \tag{3}$$

where

$F = z^T f$ = the total falsification,

$\text{MUF}_u = z^T o$ = the (unfalsified) MUF that would be reported if F were equal to 0, and

$$\begin{aligned} z^T &= [z(BI)^T \mid z(R)^T \mid z(S)^T \mid z(EI)^T] \\ &= [+1(BI)^T \mid +1(R)^T \mid -1(S)^T \mid -1(EI)^T] \ , \end{aligned} \tag{4}$$

1(BI) denotes the N(BI)-component vector with all components equal to one, and 1(R), 1(S), and 1(EI) are defined similarly.

Suppose for the moment that the inspectorate independently measured all items in the reported MUF. (Aside: assumed independence does not simplify the mathematics, but it does lead to cleaner interpretation of results.) Then the resulting data would be represented by the vector i, where

$$i^T = [i(BI)^T \mid i(R)^T \mid i(S)^T \mid i(EI)^T] \tag{5}$$

and

- i(BI) is the vector of the N(BI) inspectorate's values corresponding to items in beginning inventory,
- i(R) is the vector of the N(R) inspectorate's values corresponding to items in receipt,
- i(S) is the vector of the N(S) inspectorate's values corresponding to items in shipment, and
- i(EI) is the vector of the N(EI) inspectorate's values corresponding to items in ending inventory.

Also, let

$$\Sigma_i = \text{the covariance matrix of } i \ . \tag{6}$$

As with the operator's covariance matrix, Σ_o, the form of the inspectorate's covariance matrix Σ_i can be very complicated depending on how measurements are made.

Owing to resource constraints, the inspectorate does not measure all of the items involved. Rather, only a portion is measured. For illustration, suppose that random samples of n(BI), n(R), n(S), and n(EI) items are inspected from the beginning inventory, receipts, shipments, and ending inventory, respectively. Extensions to other (more realistic) sampling regimes are discussed later. Let $\bar{d}(BI)$, $\bar{d}(R)$, $\bar{d}(S)$, and $\bar{d}(EI)$ denote the average observed differences, operator minus inspector, for the items measured by the inspector in the beginning inventory, receipts, shipments, and ending inventory, respectively. The D statistic for this sampling procedure extrapolates these averages and is

$$\begin{aligned} D = {} & N(BI)\,\bar{d}(BI) + N(R)\,\bar{d}(R) - N(S)\,\bar{d}(S) \\ & - N(EI)\,\bar{d}(EI) \ . \end{aligned}$$

An equivalent way of writing D, and one which leads to the derivation of useful properties, is

$$D = s^T [(o + f) - i] \ , \tag{7}$$

where [(o + f) - i] represents the vector of differences that would have been observed had the inspectorate measured all N items and the vector s (for "sampling") reflects the sampling regime. In partitioned form,

$$s^T = [s(BI)^T \mid s(R)^T \mid s(S)^T \mid s(EI)^T] \ . \tag{8}$$

The j^{th} element of s(BI) is equal to

$$[s(BI)]_j = \begin{cases} N(BI)/n(BI) \text{ times the } j^{th} \text{ element of } z(BI) & \text{if the } j^{th} \text{ item in beginning inventory is measured by the inspectorate, and} \\ 0 & \text{if the } j^{th} \text{ item is not measured by the inspectorate.} \end{cases}$$

The vectors s(R), s(S), and s(EI) are defined similarly.

Note that the vector s in Eq. (8) is random—it depends on the sampling mechanism. The expected value of s is equal to z, written

$$E_s(s) = z \ , \tag{9}$$

where E_S denotes expectation with respect to the sampling mechanism and z is as defined in Eq. (4). Loosely speaking, Eq. (9) implies that each item is given weight one, on average, by the sampling, and this phenomenon is at the heart of the unbiasedness of the D statistic in evaluating falsification.

Other properties of s as defined above are found from sampling theory. Consider the covariance matrix of $s(BI)$, denoted $\Sigma_{s(BI)}$. That matrix is

$$\Sigma_{s(BI)} = k(BI)\left[I(BI) - \frac{1(BI)1(BI)^T}{N(BI)}\right] \quad (10)$$

for the constant

$$k(BI) = \frac{N(BI)[N(BI) - n(BI)]}{n(BI)[N(BI) - 1]} \quad \text{if } N(BI) > 1$$
$$= 0 \quad \text{if } N(BI) = 1$$

and $I(BI)$ the identity matrix with dimension $N(BI)$.

Because the inspectorate pursues stratified sampling with four random samples (for BI, R, S, and EI, respectively), the covariance matrix Σ_s for s has the block diagonal form

$$\Sigma_s = \begin{matrix} \Sigma_{s(BI)} & 0 & 0 & 0 \\ 0 & \Sigma_{s(R)} & 0 & 0 \\ 0 & 0 & \Sigma_{s(S)} & 0 \\ 0 & 0 & 0 & \Sigma_{s(EI)} \end{matrix} \quad . \quad (11)$$

where each of $\Sigma_{s(BI)}$, $\Sigma_{s(R)}$, $\Sigma_{s(S)}$, and $\Sigma_{s(EI)}$ have the structure as indicated in Eq. (10). More refined stratification, such as when the items in beginning inventory are themselves stratified and randomly sampled, generate other block diagonal covariance structures.

B. Mean and Variance of D

Equations (1)-(11) lay the groundwork for obtaining properties of the D statistic. Returning to Eq. (7), the expected value of D is

$$E(D) = E[s^T(o + f - i)] \; ,$$

and mutual independence of (s, o, i) implies this is

$$= F \; , \quad (12)$$

where F is the total falsification. This result depends on unbiased measurements and on a sampling regime for which $E_S(s) = z$. Many sampling regimes, such as those based on more refined stratification or on clustering, can satisfy this condition; no additional assumptions are required.

The variance of D can also be found following from Eq. (7), though the mathematics are somewhat complicated. Begin with

$$\mathrm{Var}(D) = \mathrm{Var}[s^T(o + f - i)] \; .$$

Independence of s, o, and i implies that this is equal to

$$= z^T \Sigma_o z + z^T \Sigma_i z + f^T \Sigma_s f$$
$$+ \mathrm{tr}[\Sigma_s][\Sigma_o] + \mathrm{tr}[\Sigma_s][\Sigma_i] \; . \quad (13)$$

where "tr" denotes the trace of a matrix (i.e., the sum of its diagonal elements).

C. Variance of MUF-D in the Conventional Framework

Note that the covariance between MUF and D is

$$\mathrm{Cov}(MUF,D) =$$
$$\mathrm{Cov}[z^T(o + f) \; , \; s^T(o + f - i)] \; .$$

Independence of s, o, and i leads, proceeding by conditioning and unconditioning, to

$$= \mathrm{Cov}[z^T(o + f) \; , \; s^T(o + f)]$$
$$= \mathrm{Cov}[z^T(o + f) \; , \; z^T(o + f)]$$
$$= \mathrm{Var}(MUF) \; ,$$

the variance of the operator's reported MUF. Thus, MUF-D has variance

$$\mathrm{Var}(MUF\text{-}D) = \mathrm{Var}(MUF) + \mathrm{Var}(D)$$
$$- 2\,\mathrm{Cov}(MUF,D)$$
$$= \mathrm{Var}(D) - \mathrm{Var}(MUF) \; . \quad (14)$$

This relation is illustrated, for restrictive conditions, on p. 173, Part F of the IAEA technical manual. However, it is clear from Eq. (13) that the relation is a general one that does not require facility's and/or inspectorate's measurements to be independent between strata or to have the so-called systematic/random error structure, nor does it strictly require stratified sampling by the inspectorate. (Aside. Per statistical convention, stratification is defined around the sampling mechanism instead of, for example, being defined around the physical characteristics of the items being measured. In the present illustration, there are four strata, corresponding to beginning inventory, receipts, shipments, and ending inventory.)

For any sampling regime and any (independent) inspection of reported accountability values, the variance of the MUF-D statistic can be written as Var (D) - Var (MUF) and, in more intuitive form, as

$$\begin{aligned}
\mathrm{Var\ (MUF{-}D)} &= z^T\, \Sigma_i\, z \\
&\quad + (\mathrm{tr}\,[\Sigma_s]\,[\Sigma_o] + \mathrm{tr}\,[\Sigma_s]\,[\Sigma_i]) \\
&\quad + f^T\, \Sigma_s\, f \\
&= \text{Var (inspectorate's MUF given 100\% sampling)} \\
&\quad + \text{Var (sampling : measurement errors)} \qquad (15) \\
&\quad + \text{Var (sampling : falsifications).}
\end{aligned}$$

The first term on the right hand side of equality (15), $z^T \Sigma_i z$, is the lower bound for Var (MUF-D) and represents the minimum achievable based on the quality of the measurements involved. That bound is attained for 100% sampling by the inspectorate, and MUF-D is simply the inspectorate's MUF. In that the inspectorate's measurements may be of poorer quality than those of the facility, obvious problems exist[4] if a small variance is required.

The second term,

$$(\mathrm{tr}\,[\Sigma_s]\,[\Sigma_o] + \mathrm{tr}\,[\Sigma_s]\,[\Sigma_i])$$

quantifies something of a penalty from the sampling imposed by resource constraints; when no facility measurements have been falsified, this term reflects the additional variability incurred beyond the minimum. For many common sampling mechanisms (such as stratified sampling) and error models (such as those involving so-called systematic and random error components), the second term can have a very simple form. Such form follows from the orthogonality of Σ_s from the sampling mechanism to the intraclass correlation structure induced by the error model, which is discussed in the next section.

The third term in (15), $f^T \Sigma_s f$, reflects the interaction between the sampling plan and the falsification scenario. It equals zero for special cases (e.g., no falsification or 100% inspection) but is usually positive when falsification occurs. This term also has interpretable structure for stratified sampling, as discussed in Sec. II.E, and has implications for approximating detection sensitivity, as discussed in Sec. II.F.

D. Special Error Models and Stratified Random Sampling

Following from Eq. (15), under no falsification

$$\mathrm{Var\ (MUF{-}D)} = z^T\, \Sigma_i\, z + \mathrm{tr}\,[\Sigma_s]\,[\Sigma_o] + \mathrm{tr}\,[\Sigma_s]\,[\Sigma_i]\ .$$

Note that only the second term, $\mathrm{tr}\,[\Sigma_s]\,[\Sigma_o]$, involves the facility's measurement errors. Following from Eq. (11),

$$\begin{aligned}
\mathrm{tr}\,[\Sigma_s]\,[\Sigma_o] &= \mathrm{tr}\,[\Sigma_{s(BI)}]\,[\Sigma_{o(BI)}] \\
&\quad + \mathrm{tr}\,[\Sigma_{s(R)}]\,[\Sigma_{o(R)}] \\
&\quad + \mathrm{tr}\,[\Sigma_{s(S)}]\,[\Sigma_{o(S)}] \\
&\quad + \mathrm{tr}\,[\Sigma_{s(EI)}]\,[\Sigma_{o(EI)}]\ ,
\end{aligned}$$

an equality which holds because the sampling mechanisms for BI, R, S, and EI are independent. Measurement correlations between strata do not affect this term; only those within each stratum are relevant.

Consider the first term,

$$\mathrm{tr}\,[\Sigma_{s(BI)}]\,[\Sigma_{o(BI)}]\ ,$$

the remaining terms behave similarly. Using Eq. (10),

$$\begin{aligned}
&\mathrm{tr}\,[\Sigma_{s(BI)}]\,[\Sigma_{o(BI)}] \\
&\qquad = k(BI)\ \mathrm{tr}\left[I(BI) - \frac{1(BI)\,1(BI)^T}{N(S_i)}\right][\Sigma_{o(BI)}]\ .
\end{aligned}$$

If all measured values in beginning inventory have a common "systematic error" with "systematic error variance" σ_η^2, the usual additive error model implies that the covariance matrix for the facility's measurements o(BI) has the intraclass correlation structure

$$\Sigma_{o(BI)} = \sigma_\eta^2\ 1(BI)\ 1(BI)^T + \sigma_\varepsilon^2\ I(BI)\ ,$$

where σ_ε^2 is the so-called "random error variance." The above is

$$\begin{aligned}
&\mathrm{tr}\,[\Sigma_{s(BI)}]\,[\Sigma_{o(BI)}] \\
&\qquad = \sigma_\varepsilon^2\ N(BI)\left[\frac{N(BI)}{n(BI)} - 1\right]\ . \qquad (16)
\end{aligned}$$

Importantly, the term involves only σ_ε, and the facility's so-called systematic error for measured beginning inventory is "eliminated."

Such elimination does not occur when beginning inventory consists of several types of items, where each type is measured with a different instrument. In other words, types of items measured independently should not be pooled into a single stratum that is sampled as a single entity when error models correspond to the intraclass structure and σ_η terms are important. It is primarily when falsification occurs, and the $f^T \Sigma_s f$ term can dominate the $\mathrm{tr}\,[\Sigma_s]\,[\Sigma_o]$ term in Eq. (15), that pooling can be helpful to the inspectorate. This is because an intelligent falsifier can take advantage of the stratification, should such be known to him. This is especially true in cases where inspection resources are scarce, most measurements are of high quality, and gross falsifications are of concern.

As a simple example, suppose that beginning inventory consists of 100 items: 80 items of one type and 20 items of another. The facility chooses to falsify accountability values for 10 of the 100 items by large amounts. Further, the inspectorate has resources to inspect 10 of the 100 items. If all items are pooled into a single stratum for sampling, at least one falsified item is inspected with 67% probability and, given low measurement errors, detection occurs.

Next, consider the case where the items are stratified by type, with the 80 items of the first type forming one stratum and the other 20 forming another. The two strata are sampled separately. No matter how inspection resources are allocated (e.g., inspecting 5 items of each type or inspecting 7 of the first type and 3 of the second), the facility has available a strategy to reduce the chances of inspection of a falsified item below the 67% figure above. Thus, stratification can be inefficient in dealing with scenarios involving gross falsifications. As such, the subject of how to stratify should be given some thought, as has been noted.[5]

At another extreme, if the scenarios of interest are such that the falsification component $f^T \Sigma_s f$ of Var (MUF-D) is not of great concern, then there is little motivation for pooling. Measurement errors dominate, and stratification can be helpful.

E. Falsification and Stratified Random Sampling

The third term in Eq. (15) is, for Σ_s as in Eqs. (10) and (11),

$$f^T \Sigma_s f = f(BI)^T \Sigma_{s(BI)} f(BI) + f(R)^T \Sigma_{s(R)} f(R) + f(S)^T \Sigma_{s(S)} f(S) + f(EI)^T \Sigma_{s(EI)} f(EI)$$

$$= s^2_{F(BI)}\, N(BI)^2 \frac{[1 - n(BI)/N(BI)]}{n(BI)} + s^2_{F(R)}\, N(R)^2 \frac{[1 - n(R)/N(R)]}{n(R)} + s^2_{F(S)}\, N(S)^2 \frac{[1 - n(S)/N(S)]}{n(S)} + s^2_{F(EI)}\, N(EI)^2 \frac{[1 - n(EI)/N(EI)]}{n(EI)}\,, \quad (17)$$

where $s^2_{F(BI)}$ is the usual "sample variance" of the N(BI) elements of the falsification vector f(BI) and $s^2_{F(R)}$, $s^2_{F(S)}$, and $s^2_{F(EI)}$ are defined analogously. These variances coupled with the "finite population corrections," $[1 - n/N]$, represent variability due to the sampling component. If there were no measurement errors ($\Sigma_o = \Sigma_i = 0$), the value of the D statistic would depend on the sampling distribution alone.

Note that $f^T \Sigma_s f$ is easily interpreted. Using beginning inventory for illustration, $s_{F(BI)}$ is the standard deviation of the falsification amounts for the N(BI) items. This standard deviation can be viewed as measuring the departure from uniform falsification. The term $1 - n(BI)/N(BI)$ is the portion of items that go uninspected. If, for example, $s_{F(BI)}$ and $[1 - n(BI)/N(BI)]$ are large, then $\bar{d}(BI)$ can vary a great deal because of sampling. Conversely, if sampling is not a great contributor to overall variability, then the driving factor behind Var $[\bar{d}(BI)]$ is measurement error.

F. Nonnormality of the D Statistic

As is apparent from the previous sections, normally distributed measurement errors need not imply that the D statistic be normally distributed. Indeed, the falsification scenario, reflected in the vector f, and the sampling regime, reflected in the matrix Σ_s, can have great impact on the distribution of D.

For illustration, histograms of the related MUF-D, simulated on the basis of a systems study of a MOX facility, are displayed in Figs. 1-3. The first figure, for reference, shows an approximately normal distribution for D given a total falsification of 8 kg. In the second figure, an apparent skewness is visible. Despite identical conditions in Figs. 1 and 2 (same operating facility, same sampling plan for the inspectorate, same total falsification, same measurement errors, same number of simulated trials), a change in the falsification scenario and thus the $f^T \Sigma_s f$ component of Var MUF-D as per Eq. (15) has led to increased variability. Note that the range of MUF-D values observed is nearly doubled in the latter figure. Figure 3 displays a multimodal behavior, which can occur in extreme cases.

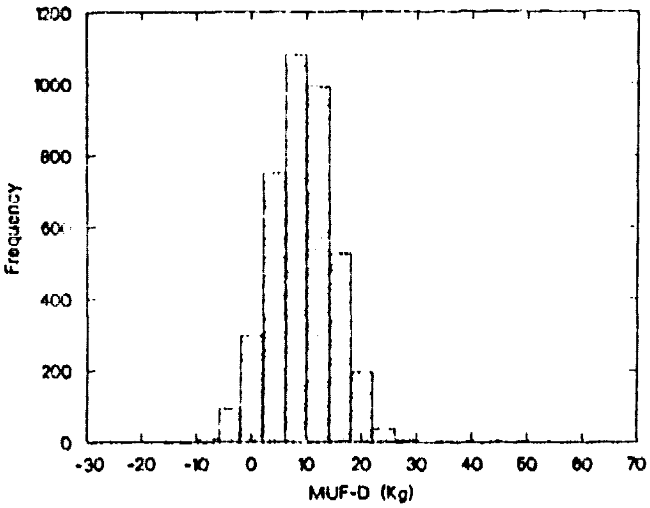

Fig. 1. Example histogram of simulated distribution for MUF-D statistic: uniform falsification totaling 8 kg.

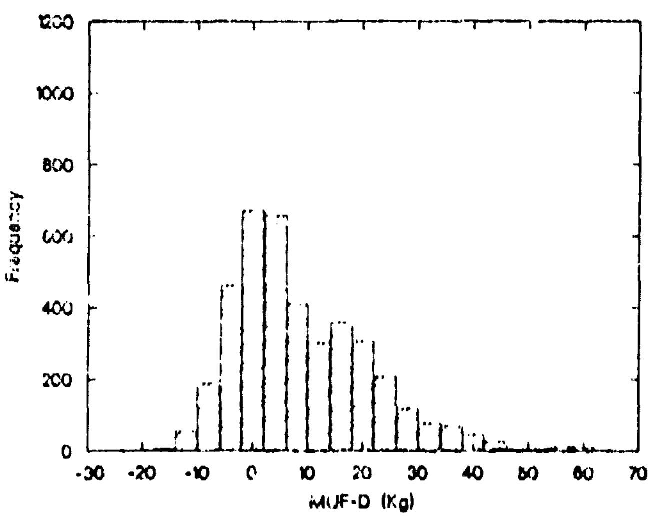

Fig. 2. Example histogram of simulated distribution for MUF-D statistic: modest falsification of a small portion of items.

Evaluation of an inspection's sensitivity must accommodate nonnormality. Although the no-falsification distribution is useful in determining threshold values for statistical tests, the detection probability, i.e., the probability that the observed D statistic will exceed a specified threshold given a particular falsification scenario, may require evaluation of the contaminated normal distribution.[6] In particular, detection probabilities derived with standard deviations appropriate for the no-falsification case used in conjunction with assumed normality can be far too optimistic.

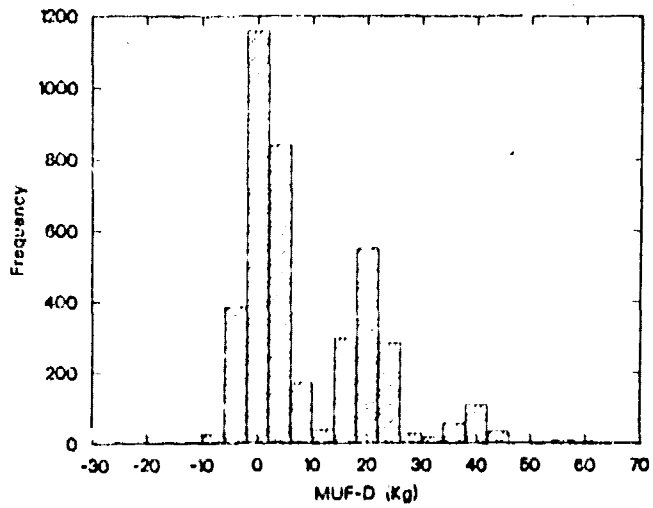

Fig. 3. Example histogram of simulated distribution for MUF-D statistic: gross falsification of very few items.

Conversely, standard deviations from worst-case scenarios can produce results that are too pessimistic.

III. THE ATTRIBUTES AND VARIABLES FRAMEWORK

A. Two D Statistics

It should be noted that some of the results of the previous section have appeared elsewhere and that a considerable literature has arisen around the conventional framework. Importantly, the formulation as in Sec. II allows for generalization to other contexts.

Consider the case where the inspectorate has the capability to measure items in any given stratum with two stratum-specific instruments. For convenience, the two types of instruments are labeled the "attributes" and "variables" instruments; in the usual setting, the variables instrument is of better quality, and cost per measurement is somewhat higher. Given an inspection that involves such measurements, use of all available information is superior to forming a D statistic based on data from variables instruments alone. Thus, the results to follow have application to data analysis as well as to planning of inspections.

Were the inspectorate to measure all items with each item's corresponding attributes and variables instrument, the resulting data would be represented by

$$a^T = [a(BI)^T \mid a(R)^T \mid a(S)^T \mid a(EI)^T] \text{ and} \qquad (18)$$
$$v^T = [v(BI)^T \mid v(R)^T \mid v(S)^T \mid v(EI)^T] ,$$

where (a,v) denotes (attributes, variables). Let Σ_a and Σ_v be the covariance matrices associated with **a** and **v**, respectively. Also let Σ_{av} denote the matrix whose $(j,k)^{th}$ element is the covariance between the j^{th} element of **a** and the k^{th} element of **v**. Many, if not most, of the elements of Σ_{av} may be zero; correlations may occur in special cases, such as when an NDA instrument is used with a short counting time for attributes purposes and a long counting time for variables purposes.

Extending the results from Sec. II.A, there are now two difference vectors, one for the inspectorate's attributes measurements and one for variables:

$$\mathbf{d(a)} = \mathbf{o} + \mathbf{f} - \mathbf{a} \quad \text{and}$$

$$\mathbf{d(v)} = \mathbf{o} + \mathbf{f} - \mathbf{v} \ . \tag{19}$$

These vectors are analogous to the vector **o + f − i** for the conventional D statistic as discussed earlier.

Owing to resource constraints, the inspector does not observe all of the elements of **d(a)** and **d(v)**. Rather, a sample is obtained. The sampling can be pursued in many different ways, such as by obtaining independent random samples[1]--one sample for the items to be monitored using the attributes instrument and another for the items to be monitored using the variables instrument. Though such a procedure is not without merit, it is somewhat inefficient because of potential redundancy. If available resources preclude extensive sampling and measurement properties are known, it is wasteful for the inspectorate to measure a given item twice (i.e., using both instruments) when the same resources could be used to inspect more items. This is especially true when gross falsifications are of concern.

For illustration, the sampling is proposed as follows, and extension to other regimes (e.g., independent random samples, approaches based on clustering or on more refined stratifications) is not difficult. The four categories of items (BI, R, S, EI) are sampled independently, as in Sec. II.A. Let n(BI) be the total number of items in the beginning inventory to be inspected, where $n(BI) = n_a(BI) + n_v(BI)$ and $[n_a(BI), n_v(BI)]$ are the number of items monitored using the (attributes,variables) instruments. A random sample of size n(BI) is obtained from the "population" of N(BI) items in the beginning inventory. Of the n(BI) sampled items, $n_a(BI)$ are selected at random and measured using the attributes instrument; the remaining $n_v(BI) = n(BI) - n_a(BI)$ items are measured using the variables instrument. Assuming that $n_a(BI)$ and $n_v(BI)$ are both greater than zero (otherwise, only one instrument is used and results are as given in Sec. II), define the sampling vectors

$\{\mathbf{s}[\mathbf{a}(BI)]\}_j = N(BI)/n_a(BI)$ if the j^{th} item is monitored by the inspectorate using the attributes instrument, and

$= 0$ else.

$\{\mathbf{s}[\mathbf{v}(BI)]\}_j = N(BI)/n_v(BI)$ if the j^{th} item is monitored by the inspectorate using the variables instrument, and

$= 0$ else.

The sampling vectors **s[a(BI)]** and **s[v(BI)]** are analogous to the vector **s(BI)** in Sec. II. Importantly, **s[a(BI)]** and **s[v(BI)]** are not independent, since no item can be inspected twice. Let **s[a(R)]** be constructed similarly to **s[a(BI)]** and so on for receipts, shipments, and ending inventory.

Define **s(a)** and **s(v)** in the anticipated way:

$$\mathbf{s(a)}^T = \{\mathbf{s}[\mathbf{a}(BI)]^T | \mathbf{s}[\mathbf{a}(R)]^T | \mathbf{s}[\mathbf{a}(S)]^T | \ \mathbf{s}[\mathbf{a}(EI)]^T\}$$

and $\tag{20}$

$$\mathbf{s(v)}^T = \{\mathbf{s}[\mathbf{v}(BI)]^T | \mathbf{s}[\mathbf{v}(R)]^T | \mathbf{s}[\mathbf{v}(S)]^T | \ \mathbf{s}[\mathbf{v}(EI)]^T\} \ .$$

It is not difficult to show that **s(a)** and **s(v)** have similar properties as the vector **s** described previously, i.e.,

$$E[\mathbf{s(a)}] = E[\mathbf{s(v)}] = \mathbf{z} \ ,$$

$$\Sigma_{\mathbf{s}[\mathbf{a}(BI)]} = k_a(BI) \left[I(BI) - \frac{\mathbf{1}(BI)\mathbf{1}(BI)^T}{N(BI)} \right] ,$$

$$\Sigma_{\mathbf{s}[\mathbf{v}(BI)]} = k_v(BI) \left[I(BI) - \frac{\mathbf{1}(BI)\mathbf{1}(BI)^T}{N(BI)} \right] ,$$

$$k_a(BI) = \frac{\{N(BI)\ [N(BI) - n_a(BI)]\}}{\{n_a(BI)\ [N(BI) - 1]\}} ; N(BI) > 1 \ ,$$

$$k_v(BI) = \frac{\{N(BI)\ [N(BI) - n_v(BI)]\}}{\{n_v(BI)\ [N(BI) - 1]\}} ; N(BI) > 1 \ ,$$

and so on for the other strata. The matrix $\Sigma_{sav(BI)}$, whose $(j,k)^{th}$ element is the covariance between $\{\mathbf{s}[\mathbf{a}(BI)]\}_j$ and $\{\mathbf{s}[\mathbf{v}(BI)]\}_k$, is

$$\Sigma_{sav(BI)} = \left[\frac{-N(BI)}{N(BI) - 1} \right] \left[I(BI) - \frac{\mathbf{1}(BI)\mathbf{1}(BI)^T}{N(BI)} \right] . \tag{21}$$

Of course, other sampling regimes generate other matrices and the concept of nonoverlapping random samples can be extended to more refined stratification. If independent random samples were obtained for the attributes and variables instruments used for the beginning inventory, the $\Sigma_s[a(BI)]$ and $\Sigma_s[v(BI)]$ would be the same as above, but $\Sigma_{sav}(BI)$ would equal the zero matrix.

The inspector can form two D statistics for each stratum. Using beginning inventory as an example, the two statistics are D[a(BI)],

$$D[a(BI)] = s[a(BI)]^T d[a(BI)] \ , \tag{22}$$

and D[v(BI)],

$$D[v(BI)] = s[v(BI)]^T d[v(BI)] \ . \tag{23}$$

B. A Generalized D Statistic

By arguments similar to those in Sec. II, it can be shown that

$$E\{D[a(BI)]\} = E\{D[v(BI)]\} = 1(BI)^T f(BI) \ ,$$

which is the overall falsification for the beginning inventory. As before, this result applies to any sampling mechanisms such that $E\{s[a(BI)]\} = E\{s[v(BI)]\} = s(BI)$ coupled with unbiased measurements and independent verification. The variances of D[a(BI)] and D[v(BI)] for beginning inventory have the same generic structure as Eq. (13), and the same can be said for D[-]'s of receipts, shipments, and ending inventory.

Because D[a(BI)] and D[v(BI)] are estimating the same thing (namely, overall falsification for items in beginning inventory), these two statistics can be combined to provide a better estimate than either one alone. The underlying principle is similar to that for a weighted average, except that the determination of weights here is not straightforward.

In the present illustration with nonoverlapping random samples, the covariance between D[a(BI)] and D[v(BI)] is, for symmetric $\Sigma_{av}(BI)$,[7]

$$\begin{aligned} \mathrm{Cov}\ \{D[a(BI)],\ D[v(BI)]\} \\ &= f(BI)^T\ \Sigma_{sav}(BI)\ f(BI) \\ &\quad + 1(BI)^T\ [\Sigma_o(BI) + \Sigma_{av}(BI)]\ 1(BI) \\ &\quad + \mathrm{tr}\ [\Sigma_o(BI) + \Sigma_{av}(BI)][\Sigma_{sav}(BI)] \ . \end{aligned} \tag{24}$$

Note that this covariance depends on the falsification vector f(BI). Unfortunately, the falsification scenario is unknown in real applications, as the value of virtually any item could be falsified to help disguise diversion of material elsewhere in the facility. Dependence of variances and covariances on f(BI) means that

(1) selection of a "good" sampling regime for inspection depends on the underlying falsification scenario, and
(2) given a sampling plan, choice of weights in a weighted average of D[a(BI)] and D[v(BI)] also depends on the underlying falsification scenario.

Were a specific scenario of interest, the variances of D[a(BI)] and D[v(BI)] and the covariance between them can be used to determine the weights in the usual weighted average to estimate overall falsification. Similar weighted averages could be constructed for all strata, and all of such estimates combined for estimation of F. The end result can be termed a generalized D statistic, generalizing from the conventional framework to the attributes/variables framework. In other words, the generalized D statistic D_G can be written

$$\begin{aligned} D_G = {} & w_{BI}\ D[a(BI)] + (1 - w_{BI})\ D[v(BI)] \\ & + w_R\ D[a(R)] + (1 - w_R)\ D[v(R)] \\ & + w_S\ D[a(S)] + (1 - w_S)\ D[v(S)] \\ & + w_{EI}\ D[a(EI)] + (1 - w_{EI})\ D[v(EI)] \ , \end{aligned} \tag{25}$$

where w_{BI} denotes the weight in the weighted average of D[a(BI)] and D[v(BI)], and w_R, w_S, and w_{EI} are similarly defined.

(Aside: if the stratification scheme for attributes measurements is not identical to that for variables measurements, estimation of falsification is somewhat more complex than above.)

C. Variance of MUF-D_G

Note that the covariance between MUF and D[a(BI)] is

$$\begin{aligned} \mathrm{Cov}\ (\mathrm{MUF}, D[a(BI)]) \\ &= \mathrm{Cov}\ (s^T(o + f),\ s[a(BI)]^T d[a(BI)]) \\ &= \mathrm{Cov}\ (s^T(o + f),\ s[a(BI)]^T[o(BI) + f(BI) - a(BI)]) \\ &= \mathrm{Cov}\ (\mathrm{MUF},\ s(BI)^T[o(BI) + f(BI)]) \ . \end{aligned}$$

Similarly, the covariance between MUF and D[v(BI)] is

$$\begin{aligned} \mathrm{Cov}\ \{\mathrm{MUF}, D[v(BI)]\} \\ &= \mathrm{Cov}\ \{\mathrm{MUF},\ s(BI)^T[o(BI) + f(BI)]\} \ . \end{aligned}$$

It follows that the covariance between MUF and D_G is

$$
\begin{aligned}
\mathrm{Cov}\ (\mathrm{MUF}, D_G) &= \mathrm{Cov}\ \{\mathrm{MUF},\ w_{BI}\ D[a(BI)] + (1 - w_{BI})\ D[v(BI)] \\
&\quad + w_R\ D[a(R)] + (1 - w_R)\ D[v(R)] \\
&\quad + w_S\ D[a(S)] + (1 - w_S)\ D[v(S)] \\
&\quad + w_{EI}\ D[a(EI)] + (1 - w_{EI})\ D[v(EI)]\} \\
&= w_{BI}\ \mathrm{Cov}\ \{\mathrm{MUF},\ D[a(BI)]\}\ (1 - w_{BI})\ \mathrm{Cov}\ \{\mathrm{MUF},\ D[v(BI)]\} \\
&\quad + w_R\ \mathrm{Cov}\ \{\mathrm{MUF},\ D[a(R)]\} + (1 - w_R)\ \mathrm{Cov}\ \{\mathrm{MUF},\ D[v(R)]\} \\
&\quad + w_S\ \mathrm{Cov}\ \{\mathrm{MUF},\ D[a(S)]\} + (1 - w_S)\ \mathrm{Cov}\ \{\mathrm{MUF},\ D[v(S)]\} \\
&\quad + w_{EI}\ \mathrm{Cov}\ \{\mathrm{MUF},\ D[a(EI)]\} + (1 - w_{EI})\ \mathrm{Cov}\ \{\mathrm{MUF},\ D[v(EI)]\} \\
&= \mathrm{Cov}\ \{\mathrm{MUF},\ z(BI)^T[o(BI) + f(BI)]\} \\
&\quad + \mathrm{Cov}\ \{\mathrm{MUF},\ z(R)^T[o(R) + f(R)]\} \\
&\quad + \mathrm{Cov}\ \{\mathrm{MUF},\ z(S)^T[o(S) + f(S)]\} \\
&\quad + \mathrm{Cov}\ \{\mathrm{MUF},\ z(EI)^T[o(EI) + f(EI)]\} \\
&= \mathrm{Cov}\ \{\mathrm{MUF},\ z(BI)^T[o(BI) + f(BI)] + z(R)^T[o(R) + f(R)] \\
&\quad + z(S)^T[o(S) + f(S)] + z(EI)^T[o(EI) + f(EI)]\} \\
&= \mathrm{Cov}\ [\mathrm{MUF},\ z^T(o + f)] \\
&= \mathrm{Var}\ (\mathrm{MUF})\ ,
\end{aligned}
$$

or the variance of MUF. Thus, as for the conventional MUF-D, it follows that

$$\mathrm{Var}\ (\mathrm{MUF} - D_G) = \mathrm{Var}\ (D_G) - \mathrm{Var}\ (\mathrm{MUF})\ . \quad (26)$$

Expressing Eq. (26) as a function of Σ_o, Σ_i, and so on is an exercise in linear algebra.

IV. DISCUSSION

Results from the previous sections are useful in the planning of inspections as well as in the analysis of resulting data. Planning is clearly a difficult issue, as there is no ideal solution. Specific plans effective at deterring/detecting one type of falsification can be ineffective against types in the "opposite direction."

To illustrate, consider the case where a small number of accountability values are falsified by large amounts. The best way to counter such action is through inspection of as many items as is practical. Presence of fixed resources motivates widespread use of attributes instruments, as the improved accuracy of the more expensive variables instrument is wasted here. Conversely, falsification of a large number of accountability values by small amounts is poorly dealt with by such an approach, as measurement noise obscures small falsifications. Thus, there does not exist an inspection that is "optimal" for all falsification scenarios.

For a specific falsification scenario of interest, one method of generating candidate sampling plans is to select plans for which

(1) costs are within resource constraints (otherwise, the "solution" is 100% inspection using the best instrumentation available), and
(2) the standard deviation of the generalized D statistic is low.

Such optimization may be nontrivial, especially if issues of how to stratify or cluster are involved or if the diminishing-returns nature of the inspection's cost vs performance need to be determined. (Aside: in item 2, use of detection probability as the criterion of interest could be entertained instead of the standard deviation, although this entails more intensive calculations for operating facilities of realistic size and gains little in terms of performance.)

Repeating the above process for several falsification scenarios, plans that are robust—i.e., that have good qualities with respect to a number of scenarios—can be generated. In this way, sample sizes for measurements using attributes and variables instruments can be obtained using a single, unified criterion.

REFERENCES

1. "IAEA Safeguards Technical Manual, Part F, Statistical Concepts and Techniques, Vol. 3," IAEA-TECDOC 261, International Atomic Energy Agency document (1982).

2. S. R. SEARLE, Linear Models, John Wiley and Sons, Inc., New York (1971).

3. F. GRAYBILL, Theory and Application of the Linear Model, Wadsworth, Belmont, California (1976).

4. M. J. CANTY, "Comments on Inspection Goal Criteria for Materials Accountancy," ESARDA Bulletin, 11, 6-7 (1986).

5. M. FRANKLIN, "The Sampling Distribution of the D Statistic," Nucl. Mater. Manage., XIII, 135-142 (1984).

6. R. AVENHAUS, Material Accountability: Theory, Verification, Applications, John Wiley and Sons, Inc., New York (1977).

7. K. K. S. PILLAY, D. STIRPE, R. R. PICARD, "International Safeguards for an Automated MOX Fuel Fabrication Facility," N-4/87-377, Los Alamos National Laboratory, Safeguards Systems Group (1987).